School performance feedback systems: conceptualisation, analysis and reflection

Adrie J. Visscher* & Robert Coe**

* University of Twente, The Netherlands
** University of Durham, England

Introduction

There is an increasing trend internationally to publish and feed back information to schools and teachers on their performance. School improvement is often the main objective for this, but accountability and the promotion of parental/student school choice also play important roles. However, practical initiatives and research in this field often fail to recognise the full complexity of improving schools through school performance feedback. Relevant issues are, for example:

- is the information fed back to schools valued and understood by school staff?
- does school performance feedback enable appropriate diagnosis and remediation to take place in schools?
- does the use of the feedback generate conflicts and tensions for teachers and administrators?
- do the features of the implementation process and school organisation influence the utilisation of school performance feedback systems?
- do schools actually improve as a result of performance feedback?
- do the feedback systems have unintended effects?

We have therefore attempted here to conceptualise 'school performance feedback systems', and to analyse the factors that have contributed to their international growth. Thereafter, the characteristics of a number of sophisticated school performance feedback systems (SPFSs) that are being used around the world will be analysed, drawing on two specific examples, followed by the presentation of a framework which sets out the factors influencing the utilisation of SPFS-information and its effects.

In the final section the status of the field is discussed by portraying the evidence on the process, problems and impact of school performance feedback from a variety of international contexts. In addition, strategies for school improvement through performance feedback, and for research in this field, are formulated.

School performance feedback systems

School performance feedback systems are defined here as information systems external to schools that provide them with confidential information on their performance and functioning as a basis for school self-evaluation. Such systems have become widespread in education in many parts of the world. They share a goal of seeking to maintain and improve the quality of schools, and arise out of a belief in the power of feedback to learn, and to produce change, often accompanied by a sense of disillusionment at the lack of impact of other models of school improvement.

This definition excludes informal, self-generated feedback and separates SPFSs from *public* systems of school performance accountability and support of school choice, which have rather different aims and contents.
The content of the information on the school's *performance* or *functioning* must be taken broadly. School performance here is likely to mean some kind of contextualised measure for fair comparison, adjusted to take account of factors beyond the control of the school ('value added'). 'Performance' may also include absolute performance measures and may equally relate to non-academic outcomes of schooling (e.g. behavioural and affective). Information on the *functioning* of schools relates to organisational and school-process measures like the resources spent, the subject matter taught, the instructional methods used and the nature of school leadership.
That the feedback should provide a basis for *self-evaluation* implies that the feedback should not simply be used for self-assessment, but that, once such judgements have been made, they ideally lead to some kind of action, such as the closer investigation of where and why the school under-performs and the development of a school improvement policy.

The origins of school performance feedback systems
A number of factors seem to have contributed to the growth of more formal school performance feedback systems in many countries over the last twenty or so years.
In many western countries in the 1980s and 90s the rise of a political climate of public sector accountability can be observed. The pressure to evaluate and report on the performance of publicly funded educational institutions may not have directly led to SPFSs, but helped to create a climate in which school performance feedback was seen as more salient than previously.
Related to this rise in accountability is the trend towards decentralisation in the administration of educational systems. As a result, schools are more likely to seek information they can utilise for school quality control, i.e. some sort of SPFS.
There is, moreover, some evidence (e.g. Murdoch & Coe, 1997) that in some countries, schools' perceptions of the unfairness of the public judgements of their effectiveness were a factor in their choice to implement a confidential value added school monitoring system (see Visscher, 2001, for an overview of the drawbacks of public school performance indicators). The published school performance information included average raw achievement of a school's students, which did not adjust for relevant features of the student intake (e.g. prior achievement levels). Schools wanted more accurate and fairer data on their own performance in order to know whether – and precisely where – improvement was really needed.
Alongside these pressures to develop systems for performance feedback, the progress made in research in the twin fields of school effectiveness and school improvement has meant that more sophisticated systems can be developed. The former line of research has resulted in a knowledge base (Scheerens & Bosker, 1997) that can be utilised in developing systems to monitor the quality of schools. An example of this is the ZEBO feedback system, which will be described below.
Research on school improvement may have influenced the development of SPFSs too, as scientific activity there showed that educational change initiatives imposed upon schools were often not very successful. Innovation and success are considered much more likely if schools themselves are convinced that something needs to be changed ('ownership'). Receiving information on how your school is doing in comparison with similar schools may be a powerful way to make you aware and determined that something needs to be changed in your organisation.

Dalin (1998), McLaughlin (1998) and Miles (1998) stress the local variability of schools, implying that general, centrally developed policies and reform strategies will not lead to educational change in all schools. Schools are considered to differ so much with respect to their performance levels (and the underlying reasons for them), their innovation capacities and contextual characteristics, that change efforts should take much more account of what is called the 'power of site or place'. Smith (1998) goes a step further. He states that as practitioners know their educational practice best they should state the goals and changes to be worked on and, after extensive training, try to accomplish those. Adaptation to the user-context can then be achieved. A SPFS may a valuable tool within this perspective on school improvement, providing timely, high-quality information on how a school 'is doing' as a basis for practitioner-led improvement actions. That may help practitioners in finding problems in their schools as well as in solving them, before it is too late. An important additional effect may be that practitioners gain a better insight into how their school works and which interventions work best in their situation.

Related to the pessimism of the school improvement authors is the view of Glass (1979) who regards education as a very complex, highly uncertain and unpredictable system on which we possess only incomplete knowledge. We should not try to find eternal truths about which of several things works well in particular circumstances, as a basis for planning and manipulating education at a large distance from the teaching-learning process in schools. Instead, we need diligent monitoring of the system in which the services are highly decentralised, and the actors are flexible and can choose from options what they consider best, rather than precisely implementing a universal approach that has been developed somewhere at a higher level.

The increase in feeding back information to schools has also been influenced by the development of multi-level and value-added data-analysis models, which enable the computation of more reliable and valid information on school functioning. The availability of computerised systems for information processing has made a significant contribution to the logistics of school performance feedback (Visscher, Wild & Fung, 2001).

Evidence about feedback effects

Given the prevalence – and growth – of performance feedback in schools, it might seem reasonable to suppose that there would be a substantial body of evidence to endorse its overall beneficial effects, and, more specifically, to identify the conditions under which its impact would be optimal. However, the evidence about feedback effects is mixed, complex and not well understood. Research results indicate that feedback can be beneficial to future performance, but it can also do harm. Moreover, the relative lack of evidence derived specifically from school contexts makes it hard to predict confidently what the effects will be in any particular case. This section sets out to review briefly what is known about feedback effects, drawing on empirical evidence and theoretical understandings from education, psychology and organisational behaviour.

Several theories have made predictions about feedback effects. The first and perhaps most influential was formulated by Thorndike (1913). His 'law of effect' saw feedback acting as either reinforcement or punishment, which determines whether subsequent behaviour is encouraged or discouraged. This theory had substantial influence on feedback research, despite inconsistencies with empirical results (Kluger and DeNisi, 1996).

A later contribution to our understanding of how feedback may influence performance comes from control theory (Podsakoff & Farh, 1989), which emphasises the discrepancy between a person's performance and their internal standards. Feedback that suggests a discrepancy between these two exists ('negative' feedback) is likely to be met with attempts to remove the discrepancy, either by increasing effort or reducing the standards, while following positive feedback, goals and effort are expected to remain stable.

Related to this idea is a third theory, that of learned helplessness (Mikulincer, 1994), which also makes predictions about the effects of feedback. In particular, it predicts that repeated negative feedback can lead to a reduction in performance below what was previously achieved.

A fourth theory that offers predictions about feedback effects is Bandura's (1991) social cognitive theory. According to this, individuals' responses to feedback depend on two regulators: a self-evaluative mechanism that compares performance to internal standards or goals and a self-efficacy mechanism that judges their capacity for attaining those goals. Feedback contributes to both these mechanisms, informing the comparison in the former and providing evidence on which to judge the latter. Social cognitive theory recognises an important tension between the effects of negative and positive feedback: although negative feedback is necessary to motivate the need for improvement, without positive feedback, individuals are unlikely to believe themselves capable of achieving it.

Finally, the theory of goal setting (Locke & Latham, 1990) offers an explanation of the role of feedback in enhancing performance when challenging goals have been set. The combination of specific and difficult goals with feedback makes satisfaction contingent on high performance, so increasing task effort, focus and persistence.

However, the most comprehensive synthesis of research on feedback effects is Kluger and DeNisi's (1996) meta-analysis and feedback intervention theory (FIT). Overall, they found an effect size of 0.4, which they interpret as 'suggesting that, on average, feedback intervention (FI) has a moderate positive effect on performance' (p. 258). However, over a third of the effects in their analysis were negative and many were close to zero. As they warn, 'F[eedback] I[ntervention]s do not always increase performance and under certain conditions are detrimental to performance' (p. 275). Similar mixed results have been found in other recent reviews and meta-analyses (e.g., in Bangert-Drowns et al., 1991; Locke and Latham, 1990, Balcazar et al., 1985) – despite the resilience of the view that feedback universally enhances performance.

According to Kluger and DeNisi's (1996) feedback intervention theory, feedback effects depend on three classes of variables: the cues in the feedback message, the nature of the task performed and on situational/personality variables. In order to predict the impact of giving feedback it is important to understand the relationships between these variables and their effects; in order to optimise its impact it is necessary to know which variables can be altered and to manipulate them accordingly.

Cues in the feedback message

Cues in the feedback message may direct attention away from the task, for example by diverting it to the self, raising issues of self-efficacy or causing individuals to focus on wider self-goals or issues of self-perception. The evidence from the meta-analysis generally supports this part of the theory, finding that when feedback was designed either to discourage or praise the recipient its effect was less. Cues may also determine attributions for success or failure and so influence future performance. On the other hand, feedback that directs attention to past performance or to learning

processes (for example, corrective feedback) may help to focus attention on task goals. However, the level of specificity of feedback can have mixed effects, since if it is too specific it may direct attention below the level necessary for optimal improvement, or if too detailed, may confuse and actually impair learning.

Coe (1998a) also examined evidence about a number of feedback characteristics not considered in Kluger and DeNisi's (1996) meta-analysis. He suggests that feedback should be made to seem credible and accurate, and perceived as providing information and supporting self-determination, rather than as surveillance or control. It should seek to generate feelings of competence but not complacency. Moreover, it should encourage recipients to attribute their level of performance to the effort they have applied or to specific, alterable factors such as their choice of strategy, and so make them feel they have control over the outcomes. Feedback should seek to make people focus on the task or on their performance relative to their past achievements, and avoid directing attention to comparisons between their performance and that of others.

The practical importance of these cues is clear, since they are explicitly manipulable. Designers of feedback systems should therefore be able to use the research evidence to optimise the impact of their feedback. What is also clear, however, is that some of the recommendations from the research (such as the need to avoid the perception of surveillance and not to focus on comparisons) are somewhat at odds with the design of some SPFSs.

The nature of the task

The relationship between task characteristics and feedback effects is also complex and not well understood, and is somewhat hampered by the lack of an adequate taxonomy of the tasks in which the effects of feedback on performance have been evaluated.

The main factor found by Kluger and DeNisi (1996) to be associated with differential effectiveness was task complexity. They report that feedback had substantial effects for simple tasks, but for complex tasks feedback had essentially no effect at all. This result is particularly interesting in view of the similarly modest effects of goal setting on complex tasks (Wood et al., 1986; DeShon and Alexander, 1996) and the fact that precisely these tasks might be expected to be most relevant for drawing inferences about school performance.

Situational/personality variables

Kluger and DeNisi (1996) identify two main factors here. The first of these is the use of a goal setting intervention. They acknowledge the weight of evidence supporting the effectiveness of goal setting although the difference found in their study was relatively modest. Certainly, since goal setting is an independently manipulable intervention, it would seem that feedback effects should be maximised by ensuring that individuals have clear, specific and challenging goals related to their task performance against which the feedback can be used to measure performance.

However, Neubert (1998) found an interesting interaction between task complexity and the role of feedback in goal setting. When Neubert's analysis was restricted to complex tasks (6 effects), the difference between combining feedback with goal setting and the latter alone rose to 1.02. This suggests that although goal setting may be less effective in general for complex tasks, the role of feedback in facilitating goal setting effects may be particularly important for these tasks. However, this area is one

where findings seem somewhat mixed and hard to interpret; making confident predictions about feedback effects is likely to be even harder.

The second characteristic identified as significant in Kluger and DeNisi's FIT, and confirmed by the meta-analysis was the extent to which self-esteem was perceived to be at threat from the feedback message. In those studies where the threat was low the effects of feedback were much higher (effect size for studies in the bottom quartile of ratings of threat to self-esteem = 0.47) than in those characterised as posing high threat (effect size = 0.08). It seems likely that performance feedback given to teachers might well be seen as extremely threatening and hence, on the basis of this result, be expected to have very little positive effect on future performance.

In addition to these two main factors, a number of other situational/personality variables have been identified in the literature as significant. For example, the FIT also predicts that feedback effects will be greatest when the cognitive demands of a task are least. A review by Coe (1998a) points to further factors such as the existing level of task-motivation and the availability of other information or instruction that may help to improve performance. Although maximising these factors is likely to be most beneficial for performance, the specific effects of feedback seem to be greatest when each of these three factors is minimised. However, a review by Goodman (1998) emphasises the difference between improving performance while practising a task and true learning. She suggests that, for tasks that provide little inherent feedback, external feedback can actually interfere with feedback derived directly from the task and so inhibit long-term learning. In the context of a SPFS, the nature of the task is likely to be fixed, so this variable cannot be manipulated.

Personality variables are also important in moderating the effects of feedback on performance. Individual differences in characteristics such as self-esteem, locus of control and achievement-orientation have been shown to influence reactions to feedback (Coe, 1998a). Kluger and DeNisi (1996) view these differences in terms of the differences in the self-goals that are salient for different people (for example, those low in self esteem are particularly anxious to avoid negative stimuli). According to their FIT, feedback that resonates with these salient goals is likely to divert attention away from the task and thus debilitate performance.

Specific application to school performance feedback

Despite the generally good fit between the data and their theory, Kluger and DeNisi (1996) are somewhat cautious in drawing implications for practice. They point out that although feedback can substantially improve performance under certain conditions, if this is achieved through an increase in task-motivation, it may not be sustained after the feedback is removed. In some cases, the costs of providing continuous feedback might outweigh the benefits of improved performance. On the other hand, if the effect of the feedback is through task-learning processes, then 'the effect may create only shallow learning and interfere with more elaborate learning' (p. 278). Further research and development of the theory are required to establish whether effects are lasting and efficient.

Of most relevance to the current discussion, however, would be studies of the impact of performance feedback conducted in school contexts. Such studies include those by Fuchs and Fuchs (1986), Cohen (1980), L'Hommedieu et al., (1988, 1990), Brandsma and Edelenbos (1992, 1998), Tymms (1995, 1997a, 1997b) and Coe (1998b). All these studies have been reviewed in more detail by Coe (2002). However, it seems that very few evaluation studies have directly addressed this issue; those that have

have generally suffered from limitations that make their findings hard to interpret unequivocally.
Given the complexity of the kinds of feedback that can be given to schools about their performance, the varying contexts of school performance, and the range of ways feedback can be provided, it is extremely difficult to make any kind of generalised predictions about its likely effects. The low level of our theoretical understandings of feedback, the lack of evidence about many aspects of its effects and, most crucially, the limitations of the evidence derived from school performance contexts make any such predictions extremely speculative. Even for a specific case in which the moderating variables referred to above are known, it would often be hard to make an unequivocal prediction; different factors often work in opposite directions and the balance of effects is usually uncertain. Moreover, such a well-described case would, in practice, be the exception rather than the rule. In short, we cannot confidently say what the benefits of giving schools performance feedback may be, or how those potential benefits may be maximised.

The factors that matter
Having examined the theoretical and empirical evidence underlying the use of school performance feedback systems, we now turn to presenting a conceptual framework for analysing the significant features of such systems. Since there is no generally accepted existing framework, the groups of factors that are believed to matter have been identified by reviewing the relevant literature in the fields of educational innovation, educational management, business administration, and computer science.
Figure 1 below presents the results of the theoretical analysis: a model depicting the assumed relationships between four groups of factors (Blocks A - D) on the one hand, and the use (Block E) and impact (Block F) of SPFSs on the other. The Figure shows that the nature and intensity of SPFS use is supposed to be influenced by the SPFS features, which in turn result from its design process. The nature of the implementation process and the characteristics of schools are also expected to influence SPFS use. The implementation process can promote SPFS use directly (e.g. by supporting schools in accomplishing the innovation), or indirectly (e.g. via training school staff in the required SPFS skills). Finally, the degree of SPFS use, and the way in which it is used, is expected to lead to both intended and unintended effects.
It is important to stress that Figure 1 is meant to clarify which factors influence SPFS use and the resultant effects, so Blocks E and F are crucial. In other words, the Figure neither shows how all factors contribute to the effects in Block F nor how other blocks in the Figure are related. If the latter had been the case, arrows between other blocks could also have been drawn.

Figure 1 also indicates that the school environment plays a role. For example, the extent to which the school board, district and the community play an active role in running schools and demand high quality may influence to which degree schools use a SPFS to improve performance. If the quality of school functioning is a hot issue, for instance where there are published league tables and 'punishments' for under-performing schools, then schools may be more inclined to improve than when external quality control is weak, and parents are unable to choose the school of their choice. The educational system can also play a more supporting role by providing schools with the resources required for change and improvement.
Each of the Blocks in Figure 1 will now be discussed more in detail. It is however impossible to provide much detail here on the reasons for selecting each factor. For more details on the backgrounds of the factors the reader can refer to Visscher (2002).

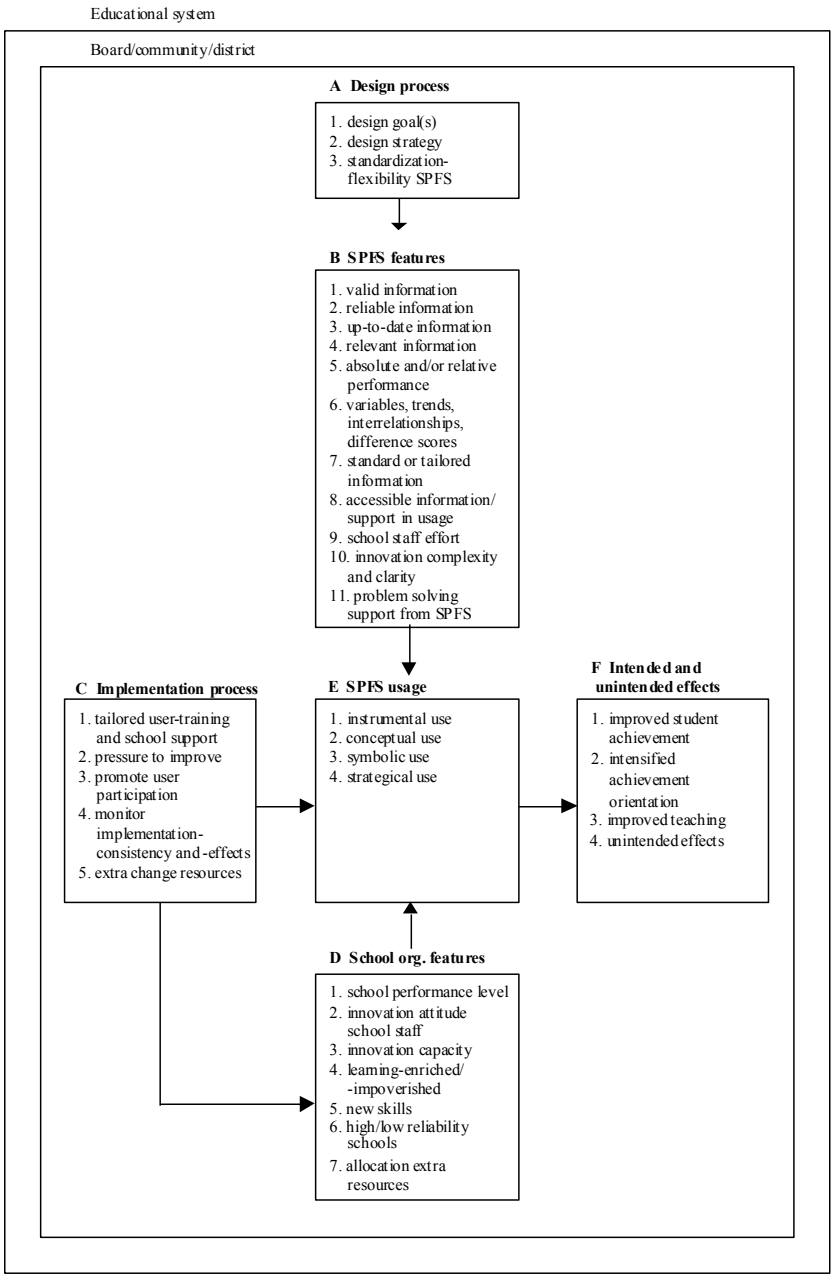

Figure 1. The relationships between the factors influencing SPFS use and effects

Block A: The design process
The process of designing a SPFS can differ in three respects.

1. The goal(s) of designing a SPFS
Although *school improvement* is the central goal for designing SPFSs, some SPFS designers may want to design systems that also serve *accountability*, and the *support of parental school choice*. Each of the three goals requires its own set of performance indicators. Whitford and Jones (1998) state that for school improvement the feedback

information should be as detailed as the complexity of schooling. SPFS designers should be clear which goal(s) they aim to serve.

2. Design strategies

Maslowski and Visscher (1999a) make a distinction between several design approaches and point to a number of relevant aspects of design strategies:

- how the problem analysis is carried out, and how objectives and means are formulated in SPFS design;
- the extent to which SPFS prototypes are developed and evaluated (formative evaluation is of great importance for design optimisation);
- the degree to which stakeholders communicate with each other and influence decisions on the desired SPFS. If practitioners have more of a say, they may develop ownership, take its findings seriously and make more effort to apply the results (cf. Huberman, 1987; Weiss, 1998);
- the creative, non-linear side of SPFS design activities.

3. The standardisation-flexibility problem

Ideally, a SPFS is so flexible that it satisfies requirements uniform to all schools as well as varying information needs among schools. In practice it will be hard to fulfil both goals completely, implying that a compromise between both is usually required.

Block B: SPFS features

SPFSs can differ in the extent to which:

1. the information fed back to schools is *valid,* e.g. value-added data versus raw data, based on multi-level analysis, or aggregated data and the extent to which SPFS data *cover school quality* (e.g. indicators on the performance of the overall school, school divisions, departments, and teachers). Fitz-Gibbon (2002) provides a typology of indicators that can support school improvement;
2. information is *reliable*;
3. information is *up-to-date*: "Timeliness is a useful feature of a report dedicated to utility" (Weiss, 1998);
4. data is *relevant* for practitioners, fits with their needs and reduces their uncertainty;
5. SPFS data indicates both *relative* and *absolute* school performance;
6. data shows values for such *factors* as *trends* over time, *relationships* between data and *differences* between *scores* measured at two or more instances (the latter for example to evaluate the effect of school policy);
7. the SPFS provides *standard* information, and allows in-depth analysis and *tailored* information for users;
8. data is presented in an *accessible* and appealing manner (e.g. quantitatively, graphically) and users are supported in *using* performance *data correctly,* e.g. the correct interpretation of value-added data;
9. the SPFS *requires* the investment of *time and effort* from school staff as a result of data collection and feedback;
10. innovation is *complex* yet *clear*, i.e. the difficulty of the required change and the success of accomplishing it for those who need to make the change.
11. the SPFS provides *user-support in problem-solving* (e.g. via a manual, computer-aided help or help line).

In Block B the importance of sophisticated SPFSs 'spitting out' high quality information is stressed. In Blocks C and D other critical success factors are stressed since SPFS quality is a necessary but insufficient precondition for the use and intended effects of SPFS. Features of schools and the change process itself are strongly related to positive outcomes of school improvement efforts (McLaughlin, 1990).

Block C: Implementation process features
Based on our review of the educational innovation literature, the following implementation process features appear to be relevant for successfully introducing SPFSs into schools:

1. A lengthy, intensive tailored *reform strategy* and *support*, e.g. assisting in school diagnosis, designing school change policies at class and school level. External change facilitators combined with information exchange via school networks: good examples can fulfil an important role here. The extent, method and content of *user-training* is very important: clarification of innovation goals and means, motivating users for innovation, developing new organisational roles, values, information processing and school improvement skills.
2. The *pressure* to improve via external targets and control, competition between schools and incentives.
3. The encouragement of *user participation* and ownership in implementation.
4. Monitoring the *consistency* of *implementing* SPFSs as well as the immediate *effects* of SPFS implementation on classrooms and student achievement. Often the implementation process needs to be adapted to the local conditions.
5. The provision of *extra innovation resources* e.g. for releasing school staff from routine work.

Block D: School organisational features
The following organisational features of schools are considered important for using a SPFS:

1. The *level of performance* of schools: relatively low levels combined with the pressure strategy may motivate schools more to try to improve performance by using a SPFS.
2. The *innovation attitude* of school staff: receptive or not.
3. The *innovation capacity*: being aware of the school's performance level, structure, culture, problems and the capacity to evaluate, to design reform goals and means, interventions at school and classroom level, experimenting, evaluating, adaptations and improving.
4. The degree to which schools promote *organisational learning*: encouragement and support via shared responsibilities for school goals/outcomes, collaborative work structures, and exchange of information, experimentation and innovation.
5. *New skills*: interpretation of SPFS output.
6. *High/low reliability schools*: the degree to which classroom and other school activities are co-ordinated.
7. Allocation of *school resources* to innovation activities.

Block E: SPFS use
What does SPFS use actually encompass? One element of use concerns the analysis and interpretation of the information received. This may not always be easy as some of the outcomes are the product of complex statistical techniques. Their correct

interpretation requires some knowledge of statistical concepts like value-added scores, correlations, and confidence intervals. Ideally, users would have been trained in this respect.

Another aspect of SPFS use concerns the utilisation of the information schools receive for improving their functioning, i.e. deciding to act to improve, and acting on it as much as possible. In the evaluation literature a distinction is made between three types of utilisation (Rossi & Freeman, 1993):

1. *direct or instrumental*: the decision-maker analyses the information before taking a decision, and bases decisions and actions on this;
2. *conceptual*: less visible but also important is the extent to which the evaluative information influences the thinking of decision-makers and as such may have an impact on their actions;
3. *convincing (symbolic)*: this type of use concerns using information in support of someone's own viewpoints in a discussion with others. Information is then used selectively to legitimise an opinion already held.

Visscher (2001) refers to Smith (1995) who presents a profound analysis of the unintended, *strategical* consequences of publishing performance data on public sector institutions. In the context of schools, the following strategic actions come to mind:

- concentrating on those students where most 'profit' can be gained;
- selective student admissions;
- removing 'difficult' students;
- teaching to the test and so concentrating on the indicators to the exclusion of other qualifications;
- consciously depressing baseline test scores to obtain high value-added scores.

Block F: The intended and unintended effects

In our view the ultimate goal of introducing SPFSs should be *improving school performance*, e.g. higher, value-added school performance scores. Proving this type of progress unequivocally will probably take a long time. In the meantime, it will be interesting to investigate to what degree a number of important prerequisites for improved school performance can be observed, such as:

1. a stronger *orientation* of school staff to *high student achievement*;
2. *improved teaching*;
3. changes in school organisational processes and structures because the use of SPFS output presupposes staff co-operation, communication and leadership.

However, because of the potential strategical use of performance indicators (Smith, 1995) it is important also to check for negative, *unintended effects* of introducing SPFSs.

The validity of the framework

Visscher's framework in Figure 1 identifies 34 factors, which emerge from a literature review as potentially influencing the success of SPFSs. The directors of five widely used SPFSs in the USA, Australia, the United Kingdom and The Netherlands were asked to use this framework to describe their projects (Teddlie et al., 2002; Hendriks et al., 2002; Gray, 2002; Rowe et al., 2002; Tymms and Albone, 2002). Based on the results of their activities, an analysis was made of the extent to which the structure and contents of Figure 1 seemed applicable to their context.

The comparison revealed that about three quarters of the factors in Figure 1 were addressed in the accounts of these SPFSs, implying that those involved explicitly noted the importance of the majority of these factors based on their experience with

developing and introducing SPFSs. This should not be too surprising since the framework is based on a review of the relevant literature, which reflects several decades of experience and research. However, as this *general* literature has been translated to the specific context of SPFSs it is good to see that the new framework applies to this particular innovation.
Nevertheless, some of the factors in the framework were not mentioned by the directors of the SPFS analysed, which is probably due to the fact that some activities were not carried out in the projects (e.g. the factor 'monitoring implementation-consistency and effects', or 'conceptual use of SPFS-information', and 'unintended effects' of SPFSs). As a result, no information is available on the extent to which these factors influence SPFS success.

The fact that so many of the factors from Visscher's framework were addressed in the accounts of the experiences with international SPFSs suggests that these factors definitely play some role. Their precise role and the degree to which they influence the success of introducing and using SPFSs are however unknown. More insight into this can only be obtained by means of empirical research; the nature of this research is explored below.

SPFS examples
In this section, two specific examples of SPFSs are described briefly, in relation to the Visscher framework and the evidence on feedback effects. One, Performance Indicators in Primary Schools (PIPS), is from the UK, the other, Self-Evaluation in Primary Education (ZEBO), from the Netherlands.

PIPS: Performance Indicators in Primary Schools (Tymms and Albone, 2002)
The stated aims of the PIPS project are concerned primarily with school improvement, centred on the belief that 'high quality information can help schools to find problems'. A wider aim, though, involves promoting a more 'scientific' approach to educational provision and attempting to 'foster a way of thinking that involves recourse to investigation and data rather than argument and opinion' (p194). The assumption seems to have been made in the design of this project that the identification of weaknesses is a key step in the process of enabling schools to improve themselves.
PIPS uses a number of different outcomes at different levels, some of which are also the key outcomes in the 'Official Accountability System' (Tymms, 1998) of school league tables, intended for accountability and to inform school choice. However, it is interesting that a concern with the quality of these statutory assessments led the project to create its own outcome measures.
The initial design of PIPS drew on an already existing model, the ALIS project, which at that time had been running successfully for nearly ten years (Fitz-Gibbon, 1996). The development of PIPS has been very much an adaptive process, absorbing contributions from schools, local education authorities, teacher and headteacher associations and, of course, researchers. This constant incorporation of formative feedback might well be seen as a validation of the expressed philosophy of the project; that receiving feedback helps one to improve. One could also interpret the attempt to evaluate the impact of the project as evidence of the authenticity of its creator's belief in 'investigation and data'.
One significant feature of PIPS is the fact that schools pay to join, either individually or as part of a group (e.g. an LEA). Schools who have paid for the information may find it easier to argue for its confidentiality and against the data being used to hold them accountable. They can choose what they do with it. Equally, the implicit

assumption in the project that simply giving schools feedback is enough to motivate them to act on it may well be true for these volunteers, but arguably might not generalise to other schools who would not choose (and pay) to join.
Flexibility is ensured by the variety of options available within the PIPS project. These include choices of year groups, assessments to be used, baselines from which to calculate value added, and levels of aggregation of the data. Some assessments are also provided in computer-adaptive form and the project supplies software for schools to analyse and interpret their data.
Concern with the technical adequacy of the measurements and models used has evidently been a key feature of PIPS. Great care has been taken with developing reliable and valid measurements. PIPS is unusual among SPFSs in not making use of multilevel models in calculating the 'value added' that is fed back to schools – a decision justified on the grounds of the greater simplicity and accessibility of ordinary (least squares) regression, and the repeated finding that residuals from the two methods differ only trivially (Fitz-Gibbon, 1997).
The format of the feedback sent to schools makes extensive use of graphical representations and though some of the information is quite complex, the PIPS staff has sought to make it readily interpretable.
Considerable evidence about the use of the data has been collected by the PIPS project. Most teachers seem to use the data to determine the strengths and weaknesses of students and classes, and to inform other teachers about this. Around half of them use the feedback to review the curriculum and to set targets. However, we do not know to what extent the data are used beyond the individual-teacher level. This evidence about the extent of use of the PIPS data, and the fact that most schools who receive it once continue to pay for it, testifies to its perceived relevance. Feedback is required to be – and is – returned quickly to schools and its relevance is explained by the fact that its content is largely a direct response to consumer demand.
PIPS has also made considerable efforts to evaluate rigorously its own impact on school performance, and specifically on pupils' achievements. Such efforts seem to be exceptional among SPFSs. It is somewhat disappointing, therefore, that the evidence available does not confirm more emphatically the positive impact of PIPS on pupil achievement.
In one randomised controlled trial, the invitation to join the project appeared to have no effect on achievement and even a slightly negative effect on attitudes. Actually joining was associated with slightly better achievement and more positive attitudes, though differences of this size could easily have arisen by chance. In another comparison, pupil gains were slightly greater in the schools that joined the project. However, none of these differences was statistically significant and the effect sizes are quite small (0.1 or less).

ZEBO: Self-Evaluation in Primary Education (Hendriks et al., 2002)
ZEBO arose from a strong tradition of evaluation and seems to set out to encompass both improvement and accountability goals under this heading. In many ways evaluation emerges as an aim in its own right, underpinned by the belief that teachers need to 'experience a deficiency' before change can turn into improvement. Self-evaluation is seen as including monitoring, analysis and diagnosis, as well as providing the basis for informing other audiences about the school's performance.
The project itself was the subject of extensive development, building on existing instruments in the Netherlands and going through several phases of piloting and improvement. The choice of which process indicators to include was partly

determined by users (principals and teachers) after the developers were unable to find full agreement within the school effectiveness research literature.
The ZEBO project was motivated in part by a concern about the technical adequacy of existing school evaluation instruments, and the developers have provided a considerable level of detail in explaining the methodology to establish the reliability and validity of the instruments used. Clearly, a high priority for the development of this SPFS was that it should meet demanding technical standards.
The efforts made to justify the validity of the feedback are particularly notable. Teachers evidently found the feedback generally in line with their previous perceptions, but the information helped to provide new insights – a difficult balance to strike for any SPFS. Its agreement with independent judgements made by the inspectorate or from peer observation is perhaps even more impressive.
ZEBO stresses the importance of information about school and classroom processes, and provides this mainly through the medium of quantitative data. Much of this is presented in the form of graphs comparing school values with those for the whole population. The project does not attempt to provide longitudinal analysis of trends but may be seen as more of a snapshot of current functioning. The interpretation of the data is largely left to the reader, supported by appropriate manuals and a telephone help-line. However, it is clear from teachers' responses to the feedback that the overwhelming majority have found it easy to interpret. It is also clear that the feedback was widely perceived as timely, relevant and complete.
The approach of collecting two perspectives on the same issue (for example, both teachers and pupils comment on classroom climate) is an interesting feature of this project. Although this practice was initially justified in terms of establishing the convergent validity of the constructs, it has also proved to be a valued feature of the feedback for schools. The kinds of discussions that arose if, for example, the two perspectives appeared to be in conflict were themselves found to be useful to the schools in which they were stimulated.
The ZEBO feedback seems to have contributed to schools' development plans, although plans of some kind were often in existence prior to their receiving the feedback, and most of the areas mentioned for development were not directly addressed in the feedback. It generally seems to have stimulated discussion at various levels within the school on matters such as policy and plans for improvement, and results were also shared with the school board, parents and the school counselling service. Schools that were inspected after receiving the feedback felt they were better prepared as a result and that it made the process more constructive.

Drawing up the balance sheet
The fact that thousands of schools around the world want these systems – and are even willing to pay for the support the SPFSs can offer – is rare for the outcomes of educational research and shows that school feedback systems are not just another result of academic, ivory-tower work with little value for the practice of education. SPFSs are practical, thriving entities that are meeting a genuine need in many schools and this may be due to the fact that teachers and school managers work in quite uncertain contexts. Information is seen as the key to managing this uncertainty.
However, the fact that people perceive a need for SPFSs does not necessarily mean that the 'pros' outweigh the 'cons', much less that their current formulations are the best they can be. Existing school feedback systems differ considerably in the domains monitored, the units of analysis, and in the aspects of the domains monitored (input, process, output, outcomes). These differences may partly be due to deliberate

decisions, for example, a theoretically based focus on school process indicators associated with school effectiveness. However, it is also possible that differences might disappear – or be reduced – if SPFS developers had more information on the nature and content of other SPFSs. In other words, exchanging information among SPFS developers could lead to more informed development and better SPFSs.

There seems to be a general notion that the process of SPFS development consists of creating something that is believed to be good for schools and, therefore, if introduced properly will not only be used, but will improve the quality of school functioning. Many SPFS developers have not analysed thoroughly the prerequisites of SPFS use in terms of the change in attitudes and skills of school staff. Often enormous resources are invested to produce high quality SPFSs. However, the extent of any systematic check on how schools deal with the result of all that developmental work appears to be negligible. As a consequence of the lack of evaluation of SPFS use, we also lack a basis for improving SPFSs and the process of introducing them into schools.

The previous discussion of the literature on feedback effects showed them to be extremely complex, not well understood and quite often even harmful. Given this complexity and the lack of high quality evaluations of interventions in schools involving feedback, it is very hard to say with confidence that we know what the likely effects of any SPFS are.

We offer, however, a number of somewhat tentative suggestions, based on the available literature, for optimising the beneficial effects of performance feedback.

The main requirement is that feedback should direct attention to an achievable gap between desired and actual performance. Two other aspects of the feedback also seem important, though the existing literature is less clear about these. Firstly, that it should facilitate genuine task learning. This may be achieved by ensuring the feedback is timely and relates to outcomes that are central to performance in the task, rather than superficial, but perhaps easily measurable aspects of it that might be improved without genuinely improving performance. Secondly, that it should be perceived as credible, accurate and fair; otherwise it can simply be rejected and produce no change in performance.

Most existing SPFSs do not seem to match very well with many of the clearest recommendations from the literature. In most cases, the feedback is presented as neutral information, and hence does not contain any specific cues (or at least not any that were intended). It seems likely, however, that the kinds of feedback presented in the SPFSs could well have conveyed considerable threats to recipients' self-esteem. If the effect of the feedback were to make people think 'How can I avoid looking bad here?' rather than 'How can I do this better?' then existing research suggests it would be unlikely to lead to real improvements in performance. Systems in which the SPFS was part of a wider system of accountability (e.g. Teddlie et al., 2002; Gray, 2002; Rowe et al., 2002) may be the most prone to this threat. It is also these systems, however, in which the feedback may have been most closely related to clear, specific and challenging task goals. Hence, there may be two conflicting tendencies here, making the overall effect hard to predict.

In all SPFSs the main comparison offered to schools is with other schools. For those doing worse than others this may have provided (at least implicitly) a clear goal of catching up, though again for these the potential threat may have been greater. For those shown to be performing better than average, however, it may have been less clear what their goals should be. Within the Louisiana accountability system (Teddlie et al., 2002) schools are also judged by, and given targets for, their improvement, which may have been a way of providing challenging but achievable targets for all

schools. Other systems (e.g. the English LEAs described by Gray, 2002) may well also have had explicit individual school targets, though it is not clear that the feedback specifically related to such targets.

With respect to the feedback's capacity for facilitating learning and its credibility, the SPFSs we have analysed seem to have fared better. They were all at pains to provide schools with genuine, relevant, timely information about important aspects of their functioning. For many the explicit intention was that schools should use this information to learn their own lessons about what is and is not working for them. Several of them also went to great lengths to ensure that the feedback not only was, but was perceived to be, credible, accurate and fair.

Although these variables have received less attention in the feedback literature, this may be more a reflection of the field's general concern with theory development and testing rather than evaluating practical applications, and the consequent lack of ecological validity of many of the contexts in which feedback effects have been evaluated (e.g. in laboratories). In practical terms, the credibility of the feedback and its focus on facilitating learning may be far more important than other features such as threats to self esteem and attributional cues, though what the truth is we do not really know at present. Certainly, we must evaluate the practical implications of these features, as well as those that are suggested more strongly by the literature to be important. Meanwhile, the differences in approach between existing systems and research evidence may say as much about the limitations of the latter as of the former.

Directions for future research

Despite the clear good intentions and plausible justifications on the part of the creators of school feedback systems, and the positive perceptions of their receivers, we cannot be confident that they are a positive benefit to the schools because they have not been evaluated adequately. A thorough and rigorous evaluation of the effects of varying SPFSs is urgently needed. Moreover, even if one were convinced that a SPFS had a positive impact this would not necessarily be sufficient reason to implement it. The costs of implementation – in terms of resources and time – might mean that a similar benefit could be achieved more easily. Thus, the question should not be simply 'does it improve performance?' but, just as importantly, 'how much and does the improvement justify the costs?'.

Next to the key question of whether the introduction of SPFSs leads to better student achievement levels, other research questions are relevant and deserve serious attention too. First, the occurrence of the other intended and unintended effects in Block F of Figure 1 need to be evaluated. Fulfilling the right conditions for SPFS-utilisation takes many years, which implies that improved student achievement levels may only be shown in the (very) long run. In the meantime the prerequisites for raising student achievement need to be evaluated: an intensified student achievement orientation, improved teaching strategies, and the execution of other individual actions and school policies meant to improve student achievement.

It is also of great importance to evaluate the extent to which unintended effects occur, e.g. do schools develop activities that show apparently higher student achievement levels which, however, may not be due to better teaching and student counselling but to window dressing?

In addition, other evaluation work may improve the quality of SPFSs around the world. As SPFSs differ considerably in their contents, in the support they offer and in what they demand from school staff, it would be valuable to have an independent group of evaluators evaluate the quality of alternative SPFSs in terms of the eleven

quality aspects in Block B of the Visscher framework. This may stimulate SPFS developers to improve the quality of their SPFSs.

Another way of improving the quality of SPFSs empirically may be the study of their usability among user groups, e.g. by providing alternative types of information (e.g. indicators on absolute or relative information, information on school performance trends over the years, and on associations between factors), alternative ways of information presentation (e.g. numerical or graphical), and investigating to what extent each alternative is appreciated and provides support. The variation between SPFSs implies that the specific features of SPFSs always have to be described explicitly when their use and impact are studied. When studying the extent and nature of SPFS use, the question should not simply be 'to what degree SPFSs are used, and with which effects?', but, 'Which types of SPFSs are used mostly in the desired ways, how have they been implemented, and in what types of schools are they used intensively?'

Alternative types of user support concern another valuable topic for research. User-support and training seem to be crucial prerequisites for successful SPFS implementation. Little empirical evidence, however, is available regarding empirically proven methods of support. Again, the relationship between the type of SPFS on the one hand and the type of user support offered on the other is important, as the two are probably closely interrelated.

Overall, we could summarise the state of the art by saying that there is still a good deal we do not yet know about the use and impact of SPFSs. We have detailed accounts of the features of a number of SPFSs themselves and some knowledge of the rationales behind their design. We also have good descriptions of the contexts in which they have been implemented and the actual process of their implementation. We know little, however, about the nature and extent of their use by schools. Nevertheless, the framework for analysing the implementation and impact of SPFSs as proposed by Visscher provides a basis for comparing their characteristics, and clearly points out the gaps in what has been investigated.

Given our ignorance on the effects of implementing any of these SPFSs, we could hardly describe their use as being evidence-based. Enlightened policy-makers and practitioners will demand something rather more before changing policy or practice. Nevertheless, there is a *prima facie* case for believing that SPFSs are a promising way forward. With this combination, the rational response must be to conduct more and better evaluations in order to produce a sounder evidence-base about the use and effectiveness of school performance feedback systems.

Our overall hypothesis is that SPFSs will be used more intensively and effectively when they are more in accordance with the factors that are identified in the research as significant, summarized in the Visscher framework presented in this article. For example, we expect that SPFS will be used more if:

- school staff have developed more 'ownership' of the SPFSs introduced into their school;
- SPFSs are more flexible in meeting varying information needs among schools;
- the information fed back is more valid, covers school quality better and allows more in depth analysis of data;
- the introduction of SPFSs is accompanied by comprehensive, tailored reform and support strategies;
- implementation of the SPFSs in schools is monitored more consistently;
- the schools into which SPFSs are introduced promote organisational learning, and

have a more developed innovation capacity.

References

Balcazar, F., Hopkins, B.L., & Suarez, Y. (1985). 'A critical, objective review of performance feedback'. *Journal of Organizational Behavior Management*, *7* (3/4), 65-89.

Bandura, A. (1991) Social cognitive theory of self regulation. *Organizational Behavior and Human Decision Processes*, *50*, 248-287.

Bangert-Drowns, R.L., Kulik, C.C., Kulik, J.A., & Morgan, M. (1991). 'The instructional effect of feedback in test-like events'. *Review of Educational Research*, *61*, pp213-38.

Brandsma, H.P., & Edelenbos, P. (1992). 'School improvement through systematic feedback of pupil level data at the school and classroom level: an experiment.' Paper presented at the European Conference on Educational Research, Enschede, The Netherlands, June 1992.

Brandsma, H.P., & Edelenbos, P. (1998). 'The effects of training programmes for principals and teachers in secondary education: a quasi-experiment based on educational effectiveness indicators.' Paper presented at the International Congress of School Effectiveness and School Improvement, Manchester, January 1998.

Brophy, J. & Good, T. (1974). *Teacher-pupil relationships*. New York: Holt, Rinehart & Winston.

Butler, R, (1988). Enhancing and undermining intrinsic motivations: the effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, *58*, 1-14

Coe, R. (1998a). Can feedback improve teaching? A review of the social science literature with a view to identifying the conditions under which giving feedback to teachers will result in improved performance. *Research Papers in Education*, *13* (1), 43-66.

Coe, R. (1998b). 'Feedback, Value Added and Teachers' Attitudes: Models, Theories and Experiments'. Unpublished PhD thesis, University of Durham.

Coe, R. (2002). 'Evidence on the role and impact of performance feedback in schools'. In: Visscher, A.J. & Coe, R. (Eds.). *School improvement through performance feedback*. Lisse/Abingdon/Exton/Tokyo: Swets and Zeitlinger.

Cohen, P.A. (1980). 'Effectiveness of student-rating feedback for improving college instruction: a meta-analysis of findings'. *Research in Higher Education*, *13*, 4, 321-341.

Dalin, P. (1998). Developing the twenty-first century school, a challenge to reformers. In A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (eds.), *International Handbook of Educational Change* (vol. 5, pp.1059-1073). Dordrecht/Boston/London: Kluwer Academic Publishers.

DeShon, R.P., & Alexander, R.A. (1996). Goal setting effects on implicit and explicit learning of complex tasks. *Organizational Behavior and Human Decision Processes*, *65,* 1, 18-36.

Fitz-Gibbon, C.T. (1996). *Monitoring Education: Indicators, Quality and Effectiveness*. London: Cassell.

Fitz-Gibbon, C.T. (2002). A typology of Indicators. In: Visscher, A.J. & Coe, R. (Eds.). *School improvement through performance feedback*. Lisse/Abingdon/Exton/Tokyo: Swets and Zeitlinger.

Fuchs, L.S., & Fuchs, D. (1986). 'Effects of systematic formative evaluation: a meta-analysis.' *Exceptional Children*, *53*, 199-208.

Glass, G.V. (1979). Policy for the Unpredictable (Uncertainty Research and Policy). *Educational Researcher*, *8*(9), 12-14.

Goodman, J.S. (1998). The interactive effects of task and external feedback on practice performance and learning. *Organizational Behavior and Human Decision Processes*, *76* (3), 223-252.

Gray, J. (2002). Jolts and Reactions: Two decades of feeding back information on schools' performance. In: Visscher, A.J. & Coe, R. (Eds.) (forthcoming). *School improvement through performance feedback*. Lisse/Abingdon/Exton/Tokyo: Swets and Zeitlinger.

Huberman, M. (1987). Steps towards an integrated model of research utilization. *Knowledge: Creation, Diffusion*, *Utilization, vol. 8*(4), 586-611.

Kluger, A.N., & DeNisi, A. (1996). The effects of Feedback Interventions on performance: a historical review, a meta-analysis, and a preliminary Feedback Intervention Theory. *Psychological Bulletin*, *119*, 2, 254-284.

L'Hommedieu, R., Menges, R.J., & Brinko, K.T. (1988). *The effects of student ratings feedback to college teachers: a meta-analysis and review of research*. Unpublished manuscript, Northwestern University, Centre for the teaching professions, Evanston, IL.

L'Hommedieu, R., Menges, R.J., & Brinko, K.T. (1990). 'Methodological explanations for the modest effects of feedback from student ratings'. *Journal of Educational Psychology*, *82*, 2, 232-241.

Locke, E.A., & Latham, G.P. (1990). *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice Hall.

Maslowski, R., & Visscher, A.J. (1999a). The potential of formative evaluation in program design models. In J. van den Akker, R. M. Branch, K. Gustafson N. Nieveen, & Tj. Plomp (eds.), *Design Methodology and Development Research in Education and Training.* Dordrecht: Kluwer Academic Publishers.

McColskey, W., & Leary, M.R. (1985). Differential effects of norm-referenced and self-referenced feedback on performance expectancies, attributions and motivation. *Contemporary Educational Psychology*, *10*, 275-284.

McLaughlin, M.W. (1998). Listening and learning from the field: tales of policy implementation and situated practice. In A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (eds.), *International Handbook of Educational Change* (vol. 5, pp.70-84). Dordrecht/Boston/London: Kluwer Academic Publishers.

McLaughlin, M.W. (1990). The Rand change agent study revisited; macro perspectives and micro realities, *Educational Researcher*, *19* (9), 11-16.

McPherson, R.B., Crowson, R., & Pitner, N.J. (1986). *Managing Uncertainty: administrative theory and practice in education*. Columbus: C.E. Merril Publishing Company.

Mikulincer, M. (1994). *Human learned helplessness: A coping perspective*. New York: Plenum Press.

Miles, M.B. (1998). Finding Keys to School Change: A 40-year Odyssey. In A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (eds.), *International Handbook of Educational Change* (vol. 5, pp.37-39). Dordrecht/Boston/London: Kluwer Academic Publishers.

Murdoch, K. and Coe, R. (1997). *Working with ALIS: a study of how schools and colleges are using a value added and attitude indicator system*. Durham: School of Education, University of Durham, United Kingdom.

Neubert, M.J. (1998). The value of feedback and goal setting over goal setting alone and potential moderators of this effect: a meta-analysis. *Human Performance, 11* (4) 321-335.

Podsakoff, P.M., & Farh, J.L. (1989). Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes*, *44*(1), 45-67.

Rossi, P.H., & Freeman, H.E. (1993). *Evaluation; a systematic approach.* Newbury Park/London/New Delhi: Sage.

Rowe, K.J., Turner R. & Lane, K. (2002). Performance feedback to schools of students' year 12 assessments: The VCE data project. In: Visscher, A.J. & Coe, R. (Eds.) *School improvement through performance feedback*. Lisse/Abingdon/Exton/Tokyo: Swets and Zeitlinger.

Scheerens, J. & Bosker, R.J. (1997). *The foundations of educational effectiveness*. Oxford: Elsevier Science Ltd.

Smith, P. (1995). On the Unintended Consequences of Publishing Performance Data in the Public Sector. *International Journal of Public Administration, 18* (2&3), 277 -310.

Smith, L.M. (1998). A kind of educational idealism: integrating realism and reform. In A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (eds.), *International Handbook of Educational Change* (vol. 5, pp.100-120). Dordrecht/Boston/ London: Kluwer Academic Publishers.

Teddlie, C., Kochan, S. & Taylor, D. (2002). The ABC+ model for school diagnosis, feedback, and improvement. In : Visscher, A.J. & Coe, R. (Eds.) (forthcoming). *School improvement through performance feedback*. Lisse/Abingdon/Exton/Tokyo: Swets and Zeitlinger.

Thorndike, E.L. (1913). *Educational Psychology. Volume 1: The original nature of man*. New York: Columbia University Teachers College.

Tymms P.B. (1995). 'Influencing educational practice through performance indicators.' *School Effectiveness and School Improvement*, *6*, 2, 123-145.

Tymms, P. (1997a). 'Responses of headteachers to value added and the impact of feedback' The Value Added National Project, Technical Report: Primary 3. London: SCAA.

Tymms, P. (1997b). 'The impact of indicators on primary schools.' Paper presented at Evidence-Based Policies and Indicator Systems Conference, University of Durham, July 1997.

Tymms, P.B. (1998). Accountability and Quality Assurance. In C. Richards & P. Taylor (Eds.), *How shall we school our children?* (pp. 171-181). London: Falmer Press.

Tymms, P. and Albone, S. (2002) 'Performance Indicators in Primary Schools'. In: Visscher, A.J. & Coe, R. (Eds.). *School improvement through performance feedback*. Lisse/Abingdon/Exton/Tokyo: Swets and Zeitlinger.

Van Vilsteren, C.A. & Visscher, A.J. (1987). Schoolwerkplanning: mogelijk in schoolorganisaties? [School work planning: possible in school organisations?] In B. Creemers, J. Giesbers, C. van Vilsteren & C. van der Perre (Eds.), *Handboek Schoolorganisatie en onderwijsmanagement* (pp. 6120-6124). Alphen aan den Rijn: Samson.

Visscher, A.J. (2001). Public School Performance Indicators: problems and recommendations. *Studies in Educational Evaluation, 27*(3), 199-214.

Visscher, A.J. (2002). A framework for studying school performance feedback systems. In: Visscher, A.J. & Coe, R. (Eds.) *School improvement through performance feedback*. Lisse/Abingdon/Exton/Tokyo: Swets and Zeitlinger.

Visscher, A.J. & Coe, R. (2002). Drawing up the balance sheet for school performance feedback systems. In: Visscher, A.J. & Coe, R. (Eds.) *School improvement through performance feedback*. Lisse/Abingdon/Exton/Tokyo: Swets and Zeitlinger.

Visscher, A.J., Wild, P., & Fung, A. (eds.) (2001). *Information Technology in Educational Management; synthesis of experience, research and future perspectives on computer-assisted school information systems.* Dordrecht/Boston/London: Kluwer Academic Publishers.

Weiss, C.H. (1998). Improving the use of evaluations: whose job is it anyway? In Reynolds, A.J. & Walberg, H.J. (eds.), *Advances in Educational Productivity*, volume 7, pp.263-276. Greenwich/London: JAI Press.

Whitford, B.L., & Jones, K. (1998). Assessment and accountability in Kentucky: how high stakes affects teaching and learning. In A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (eds.), *International Handbook of Educational Change* (vol. 5, pp.1163-1178). Dordrecht/ Boston/London: Kluwer Academic Publishers.

Wood, R.E., Mento, A.J., & Locke, E.A. (1987). Task complexity as a moderator of goal effects: a meta-analysis. *Journal of Applied Psychology*, *72*, 416-25.